Exponentialfördelningen

Används ofta för att beskriva livslängder

$\xi \sim Exp(\lambda)$, $\lambda > 0$, sägs vara *exponentialfördelad* om den har tätheten

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \Rightarrow F(x) = \int_0^x \lambda e^{-\lambda t}\, dt = \left[-e^{-\lambda t}\right]_0^x = 1 - e^{-\lambda x}$$

$$\mathbb{E}[\xi] = \frac{1}{\lambda}, \quad Var(\xi) = \frac{1}{\lambda^2}$$

"skev" fördelning

Ex

Servicetid $\xi \sim Exp(\lambda)$, $\mathbb{E}[\xi] = 5$, i ett bibliotek.

$\Rightarrow \mathbb{P}(\text{servicetid längre än 10 minuter}) =$

$= \mathbb{P}(\xi > 10) = 1 - F(10) = 1 - (1 - e^{-10/5}) = e^{-2} \approx 0.135$

$Var(\xi) = \frac{1}{\lambda^2} = \frac{1}{(\frac{1}{5})^2} = 25$

Ex

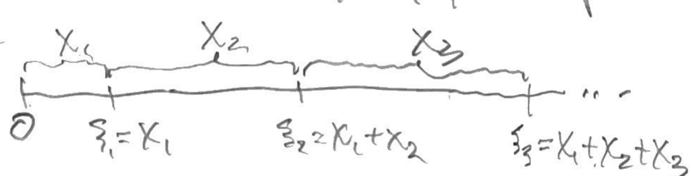
Finn $\mathbb{P}(\xi < \mathbb{E}[\xi])$, när λ okänd för $\xi \sim Exp(\lambda)$.

Lösn:

$\mathbb{P}(\xi < \mathbb{E}[\xi]) = F(\frac{1}{\lambda}) = 1 - e^{-\lambda \frac{1}{\lambda}} = 1 - e^{-1} \approx 0.633$

Poissonprocesser

Låt $X_1, X_2, \ldots$ vara oberoende och $Exp(\lambda)$-fördelade och låt $\xi_i = X_1 + \ldots + X_i$, $i = 1, 2, \ldots$ representera händelsetider



Vi kallar $N(t) = \#\{\xi_i \leq t : i = 1, 2, \ldots\}$, $t \geq 0$, dvs antalet händelser t.o.m. tiden t, för en *Poissonprocess med intensitet λ*.

Namnet kommer från att

$$\mathbb{P}(N(t)=k) = \mathbb{P}(\xi_k \le t, \xi_{k+1} > t) = \mathbb{P}(X_1+\dots+X_k \le t,\ X_1+\dots+X_{k+1} > t)$$

$$= \dots = \int_0^t \mathbb{P}(t - X_{k+1} < r \mid \xi_k = r)\, f_{\xi_k}(r)\, dr =$$

$$= \int_0^t \frac{\lambda^k}{(k-1)!} r^{k-1} e^{-\lambda r} \cdot e^{-\lambda(t-r)}\, dr = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$, dvs $N(t) \sim Poi(\lambda t)$

Poisson processer används för att modellera t.ex.
- # kunder som kommer till en affär
- # personer som ringer till en telefonväxel

Weibullfördelning

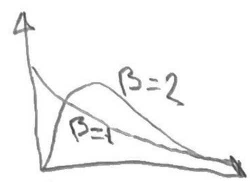
Används ofta för att beskriva tid till fel.

ξ är Weibull(α, β)-fördelad om ξ har täthet

$$f(x) = \begin{cases} \frac{\beta x^{\beta-1}}{\alpha} e^{-x^\beta/\alpha}, & x \geq 0 \\ 0, & x < 0 \end{cases} \Rightarrow F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x^\beta/\alpha}, & x \geq 0 \end{cases}$$

$$\mathbb{E}[\xi] = \begin{cases} \alpha, & \beta = 1 \\ \frac{\sqrt{\pi}\sqrt{\alpha}}{2}, & \beta = 2 \end{cases}, \quad Var(\xi) = \begin{cases} \alpha^2, & \beta = 1 \\ \alpha\left(1 - \frac{\sqrt{\pi}}{4}\right), & \beta = 2 \end{cases}$$

Notera: Weibull$(\alpha, 1) = Exp\left(\frac{1}{\alpha}\right)$



"Skev" fördelning

Ex

Man utnötningstestar lager och registrerar tiden till fel, $\xi \sim \text{Weibull}(110, 1.2)$, mätt i 10^5-cykler.

$$\mathbb{P}(\xi < 160) = 1 - e^{-160^{1.2}/110} \approx 0.9819 \quad 10^5\text{-cykler}$$

Medianen, Md, är det värde för vilket $F(Md) = 0.5$

$$\mathbb{P}(\xi < Md) = 1 - e^{-Md^{1.2}/110} = 0.5 \Rightarrow 0.5 = e^{-Md^{1.2}/110} \Rightarrow$$

$$-\frac{Md^{1.2}}{110} = \log 0.5 \Rightarrow Md \approx (-110 \log 0.5)^{1/1.2} = 37.03 \quad 10^5\text{-cykler}$$

Normalfördelningen

Detta är den viktigaste av alla fördelningar och den dyker upp på många ställen inom (den matematiska) vetenskapen.

Def

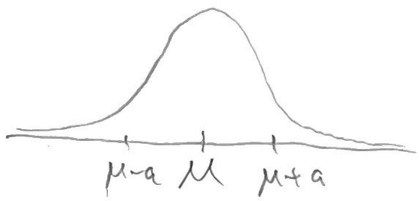
En stokastisk variabel $\xi \sim N(\mu, \sigma^2)$ (alt. $\xi \sim N(\mu, \sigma)$) kallas <u>normalfördelad/Gaussiskt fördelad</u> om dess täthet ges av

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

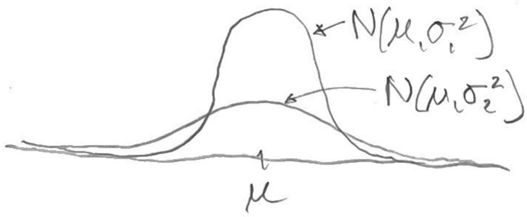
Notera att parametrarna här råkar vara ξs väntevärde och varians (alt. standardavvikelse), dvs

$$\mathbb{E}[\xi] = \mu \quad \text{och} \quad Var[\xi] = \sigma^2$$

Detta är en symmetrisk fördelning centrerad kring $\mathbb{E}[\xi]=\mu$



$$f(\mu+a)=f(\mu-a)$$



$\sigma_1<\sigma_2$ Ju högre variansen σ^2 är desto mer "tillplattad" är tätheten.

Notera att båda dessa tätheter integrerar till 1.

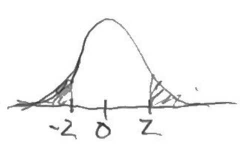
En höjning av μ resulterar i en förskjutning av tätheten (till höger)

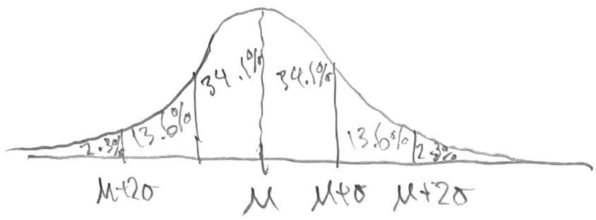
Eftersom $\frac{\xi-\mu}{\sigma}\sim N(0,1)$ så räcker det att vi känner till fördelningsfunktionen för <u>standardnormalfördelningen</u>, $N(0,1)$:

$F_\xi(a)=\mathbb{P}(\xi\le a)=\mathbb{P}\left(\frac{\xi-\mu}{\sigma}\le\frac{a-\mu}{\sigma}\right)=\Phi\left(\frac{a-\mu}{\sigma}\right)$, där $\Phi(x)=$
$=\mathbb{P}(Z\le x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-x^2/2\sigma^2}dx$ för $Z\sim N(0,1)$.

Tyvärr har $\Phi(x)$ ingen trevlig matematisk form men längst bak i boken finns en tabell för $\Phi(x)$, givet en massa värden för x.

Vill man räkna ut $\Phi(-x)$ så använder man att $\Phi(-x)\underset{\uparrow \text{pga symmetrin}}{=}1-\Phi(x)=\mathbb{P}(Z>x)$





Ex

$$\xi \sim N(3.7, 2.7^2) \Rightarrow Z = \frac{\xi-\mu}{\sigma} = \frac{\xi - 3.7}{2.7} \sim N(0,1)$$

$$\Rightarrow P(\xi < 4) = P\left(\frac{\xi-3.7}{2.7} < \underbrace{\frac{4-3.7}{2.7}}_{0,11}\right) = P(Z < 0.11) = \Phi(0.11) \approx 0.5438$$

från tabell

$$P(Z<-1.2) = \Phi(-1.2) = 1-\Phi(1.2) \underset{\text{tabell}}{=} 1-0.8849 = 0.1151$$

$$P(Z>1.7) = 1-\Phi(1.7) \underset{\text{tabell}}{=} 1-0.9554 = 0.0464$$

$$P(-1.2<Z<1.7) = P(Z<1.7)-P(Z<-1.2) = \Phi(1.7)-(1-\Phi(1.2)) =$$
$$= 0.9554 - 0.1151 = 0.8403$$

$$P(|Z|<2.3) = P(-2.3<Z<2.3) = \Phi(2.3)-(1-\Phi(2.3)) = 2\Phi(2.3)-1$$
$$= 2\cdot 0.9893 - 1 = 0.9786$$

Ex Livslängden ξ för en viss sorts glödlampor antas vara $N(\mu,\sigma^2)$-fördelad, $\mu=1700$, $\sigma=100$ (timmar).

a) Finn $P(1480 \le \xi \le 1880)$

b) I ett belysningsnät används 500 glödlampor. Hur länge bör man vänta innan 25 lampor har slocknat?

Lösn

a) $P(1480 \le \xi \le 1880) = P(\xi \le 1880) - P(\xi \le 1480) = P\left(Z \le \frac{1880-1700}{100}\right) - P\left(Z \le \frac{1480-1700}{100}\right) = \Phi(1.8) - \Phi(-2.2) = \Phi(1.8) - (1-\Phi(2.2)) = 0.9641 - (1-0.9861) = 0.9502$

b) $\frac{25}{500} = 0.05$ är andelen lampor som skall ha slocknat i genomsnitt.
Vi tolkar detta som att vi skall finna tiden t sådan att

$$0.05 = P(\xi \leq t) = \left(Z \leq \frac{t-1700}{100}\right) = \Phi\left(\frac{t-1700}{100}\right)$$

Från tabellen har vi $\Phi(-1.645) = 0.05$

$$\Rightarrow -1.645 = \frac{t-1700}{100} \Leftrightarrow t = 1700 - 164.5 =$$

$= 1535.5$ timmar.

Funktioner av stokastiska variabler

Om ξ är diskret, a och b konstanter:

$$\mathbb{E}[a\xi+b] = \sum_{x\in\Omega}(ax+b)f(x) = a\underbrace{\sum_{x\in\Omega} x f(x)}_{=\mathbb{E}[\xi]} + b\underbrace{\sum_{x\in\Omega} f(x)}_{=1} = a\mathbb{E}[\xi]+b$$

Om ξ är kontinuerlig så byter vi ut summorna mot integraler $\Rightarrow$ Räkneregeln gäller även då. Vidare,

$$\mathrm{Var}(a\xi+b) = \mathbb{E}[((a\xi+b)-\mathbb{E}[a\xi+b])^2] = \mathbb{E}[(a\xi+\cancel{b}-a\mathbb{E}[\xi]-\cancel{b})^2]$$
$$= \mathbb{E}[(a\xi-a\mathbb{E}[\xi])^2] = \mathbb{E}[a^2\xi^2-2a^2\xi\underbrace{\mathbb{E}[\xi]}_{konst}+a^2\underbrace{\mathbb{E}[\xi]^2}_{konst}] =$$
$$= a^2\mathbb{E}[\xi^2]-2a^2\mathbb{E}[\xi]\mathbb{E}[\xi]+a^2\mathbb{E}[\xi]^2 = a^2(\mathbb{E}[\xi^2]-\mathbb{E}[\xi]^2) = a^2\mathrm{Var}(\xi)$$

$\therefore$ V.v. är linjärt men ej variansen

Ex

Antal gånger som rullsand går sönder	Sannolikhet
0	0.1
1	0.48
2	0.22
3	0.2

Årsavgift på 15000 kr + 7500 kr per gång ngn kommer och lagar fel.

Finn v.v. och varians för årliga kostnaden.

Lösn.: ξ = # gånger går sönder, η = total servicekostn $= 15000+7500\xi$

$\mathbb{E}[\xi] = 0\cdot0.1+1\cdot0.48+2\cdot0.22+3\cdot0.2 = 1,52$

$\mathbb{E}[\xi^2] = 0^2\cdot0.1+1^2\cdot0.48+2^2\cdot0.22+3^2\cdot0.22 =$

$\mathrm{Var}(\xi) = \mathbb{E}[\xi^2]-\mathbb{E}[\xi]^2 = 0.8496$

$\mathbb{E}[\eta] = 15000+7500\underbrace{\mathbb{E}[\xi]}_{=1.52} = 26400$

$\mathrm{Var}(\eta) = \mathrm{Var}(15000+7500\xi) = 7500^2\overbrace{\mathrm{Var}(\xi)}^{=0.8496} = 47790000$

Om ξ är diskret och g är en reellvärd funkt.:

$$\mathbb{E}[g(\xi)] = \sum_{x\in\Omega} g(x)f(x) = \sum_{x\in\Omega} g(x)\mathbb{P}(\xi = x)$$

Om ξ istället är kontinuerlig:

$$\mathbb{E}[g(\xi)] = \int_{-\infty}^{\infty} g(x)\underbrace{f(x)}_{\text{täthet}}\,dx$$

Ex Om $\xi \sim R[0,1]$, vad är $\mathbb{E}[e^{\xi}]$?

Lösn

$g(x) = e^x$, $\mathbb{E}[e^{\xi}] = \mathbb{E}[g(\xi)] = \int g(x) f(x)\,dx =$

$= \int_0^1 e^x \frac{1}{1-0}\,dx = [e^x]_0^1 = e - 1 \approx 1.718$

Kom ihåg att händelser A och B är oberoende om $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Def

Två stokastiska variabler ξ_1 och ξ_2 är oberoende om $\mathbb{P}(\{\xi_1 \in A_1\}\cap\{\xi_2 \in A_2\}) = \mathbb{P}(\xi_1 \in A_1)\mathbb{P}(\xi_2 \in A_2)$ för alla $A_1, A_2 \subseteq \mathbb{R}$. Alternativt:

$$\mathbb{P}(\{\xi_1 \le x_1\}\cap\{\xi_2 \le x_2\}) = \mathbb{P}(\xi_1 \le x_1)\mathbb{P}(\xi_2 \le x_2) \text{ för alla } x_1, x_2 \in \mathbb{R}$$

$$\mathbb{P}(\{\xi_1 \le x_1\}\cap\{\xi_2 \le x_2\}) = \mathbb{P}(\xi_1 \le x_1, \xi_2 \le x_2)$$

Ex Fördelningarna för efterfrågan två produkter.

x	$P(\xi=x)$	y	$P(\eta=y)$
0	0.11	0	0.15
1	0.12	1	0.25
2	0.19	2	0.30
3	0.38	3	0.30
4	0.20		

ξ och η ober.

a) $P(\text{ingen produkt efterfrågas}) =$

$$= P(\xi=0, \eta=0) \overset{\text{ober}}{=} P(\xi=0)P(\eta=0) = 0.11 \cdot 0.15 = 0.0165$$

b) $P(\text{exakt två enheter av samma produkt säljs}) = P(\xi=2, \eta=0) + P(\xi=0, \eta=2) =$

$$\overset{\text{ober}}{=} P(\xi=2)P(\eta=0) + P(\xi=0)P(\eta=2) = 0.19 \cdot 0.15 + 0.11 \cdot 0.30 =$$

$$= 0.0615$$

c) $P(\text{minst 3 A-produkter och högst en B-produkt efterfrågas}) = P(\xi \geq 3, \eta \leq 1) \overset{\text{ober}}{=} P(\xi \geq 3)P(\eta \leq 1) =$

$$= (0.38 + 0.2)(0.15 + 0.25) = 0.232$$

d) Man lagerför 10 A-produkter. Varje ger en bruttovinst på 500 kr. Lagerkostnaden är 50 kr/enhet. Vad är förväntad daglig nettovinst?

$\gamma = \text{nettovinst} = 500\xi - (10-\xi)\cdot 50$

$$E[\gamma] = E[500\xi - (10-\xi)\cdot 50] = (500\cdot 0 - (10-0)\cdot 50)\overbrace{P(\xi=0)}^{=0.11} +$$

$$+ (500\cdot 1 - (10-1)\cdot 50)P(\xi=1) + \ldots + (500\cdot 4 - (10-4)\cdot 50)P(\xi=4) = 842$$

Givet stokastiska variabler $\xi_1, \dots, \xi_n$ (som ej nödvändigtvis är oberoende) så gäller

$$\mathbb{E}[a_1\xi_1 + b_1 + \dots + a_n\xi_n + b_n] = \left(\sum_{i=1}^n a_i \mathbb{E}[\xi_i]\right) + \left(\sum_{i=1}^n b_i\right)$$

för konstanter $a_1, \dots, a_n$ och $b_1, \dots, b_n$.

Vidare gäller att

$$\mathrm{Var}(a_1\xi_1 + b_1 + \dots + a_n\xi_n + b_n) = \sum_{i=1}^n a_i^2 \mathrm{Var}(\xi_i)$$

om $\xi_1, \dots, \xi_n$ är oberoende. Om de inte nödvändigtvis är oberoende lägger vi till $2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_i a_j \mathrm{Cov}(\xi_i, \xi_j)$, där kovariansen definieras som $\mathrm{Cov}(\xi, \eta) = \mathbb{E}[(\xi - \mathbb{E}[\xi])(\eta - \mathbb{E}[\eta])]$. $\mathrm{Cov}(\xi, \eta)$ är ett sätt att kvantifiera beroenden mellan ξ och η. Notera att $\mathrm{Var}(\xi) = \mathrm{Cov}(\xi, \xi)$.

<u>Ex</u> ξ är en stok. var. med $\mathbb{E}[\xi] = \mu$ och $\mathrm{Var}(\xi) = \sigma^2$. Låt Z vara "standardiseringen" av ξ, dvs $Z = \frac{\xi - \mu}{\sigma}$. Finn $\mathbb{E}[Z]$ och $\mathrm{Var}(Z)$.

$$\mathbb{E}\left[\frac{\xi - \mu}{\sigma}\right] = \frac{\mathbb{E}[\xi] - \mu}{\sigma} = 0, \quad \mathrm{Var}\left(\frac{\xi - \mu}{\sigma}\right) = \frac{\mathrm{Var}(\xi - \mu)}{\sigma^2} = \frac{\mathrm{Var}(\xi)}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1$$

~~Ex Svenska hushålls månadskostnad för mat antas normalfördelad med standardavvikelse 650 kr. Antag att man inte vill att stickprovsmedelvärdet $\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n} \xi_i$ skiljer sig från populationsmedelvärdet, dvs $\mu = \mathbb{E}[\xi]$, med mer än 50 kr. Vad bör n vara för att slh för denna händelse skall vara mindre än 0,1?~~